Development of Analytical Ultracentrifugation Standards

PHYS4995 Honours Thesis

Reece Martin[1,2]
Borries Demeler[2, 3]

[1]Department of Physics and Astronomy, University of Lethbridge, Alberta, Canada, T1K 3M4.
[2]Department of Chemistry and Biochemistry, University of Lethbridge, Alberta, Canada, T1K 3M4.
[3]Department of Chemistry and Biochemistry, University of Montana, Missoula, MT, USA, 59812.

Submission Date: 21/4/23

Contents

1 Abstract

This project entails the hydrodynamic study of biological macromolecules in the solution phase, specifically on double-stranded DNA molecules as a biological system that can be purified and prepared in colloidal, aqueous solutions suitable for study by analytical ultracentrifugation (AUC). The goal is to produce, purify, and analyze by AUC, preparations of double-stranded DNA molecules of defined length, intended for use as AUC standards for downstream applications in the biopharma industry to validate AUC results. The following sections will discuss the processes by which double-stranded DNA samples are produced, purified, and prepared for use in the AUC, and how these processes have been optimized. As such, once samples have been prepared, they are placed in an analytical ultracentrifuge under varying conditions, such as different rotor speeds, different temperatures, and in buffers with different ionic strengths. The results are then analyzed using the UltraScan software, which uses remote supercomputers to derive hydrodynamic parameters such as the sedimentation coefficient, the diffusion coefficient, and the partial concentrations of each solute in the mixture. From the hydrodynamic parameters, the partial specific volume and the anisotropy of the components in the mixtures can be derived. The data acquisition, processing, and refinement techniques used to derive the hydrodynamic parameters and extract meaningful results from them will be detailed as well. At the current stage of this project, the goal is to achieve batch-to-batch reproducibility in the generation and purification of double-stranded DNA fragments. Finally, the long-term expected outcome of this project is the development of Good Manufacturing Practice molecular standards for analytical ultracentrifugation.

2 Introduction

The goal of this project is to develop Good Manufacturing Practice (GMP) molecular standards for analytical ultracentrifugation (AUC) using double-stranded DNA molecules. AUC is a first-principles biophysical characterization technique that can used to determine the physical properties of macromolecules in the solution phase. This is achieved by analyzing the behaviour of the macromolecules while sedimenting under the influence of strong centrifugal fields (Cole et al., 2008). Molecular standards are prevalent across the biopharmaceutical industry and serve in the upkeep of the strict quality control standards that are mandated in GMP programs orchestrated by numerous governing bodies across the globe (Savelyev et al., 2020). With the increasing prevalence of AUC across the biopharmaceutical and biotechnology industries the precise calibration of AUC instruments can vary from lab to lab in a statistically-significant manner. Molecular standards allow for relative calibration differences between instruments to be characterized and corrected for (Savelyev et al., 2020). As a result, molecular standards will serve an important role in AUC data and instrument operation validation within the GMP space. These standards will also aid in testing and validating the data fitting methods of the UltraScan analysis software in terms of the reproducibility in the methods used and individual AUC instruments. The molecular standards will be validated by multiple orthogonal metrics including dynamic light scattering (DLS), and comparisons to previously-validated National Institute of Standards and Technology (NIST) reference materials. Double-stranded DNA molecules are suitable as molecular standards for AUC as their molar masses are known as the sequence length and composition can be precisely controlled using recombinant techniques. DNA molecules also provide the advantage of being highly stable in certain buffer conditions (Matange et al., 2021). The following sections will explore how these double-stranded DNA molecules are produced, purified, and analyzed by AUC.

Figure 1: Laboratory space at the Canadian Center for Hydrodynamics.

3 Rationale for Molecular Standards

The primary objective of this project was to develop double-stranded DNA molecular standards for analytical ultracentrifugation (AUC). Analytical ultracentrifugation is a first-principles biophysical characterization technique. The instrument applies large centrifugal fields to sediment macromolecules in solution, and uses various optical detection methods to provide detailed information on the macromolecules based upon their sedimenting behaviour (Cole et al., 2008). Since AUC is a first-principles technique, reference materials are not needed, provided the instrument's rotor speed, optical magnification, and rotor temperature are correct (Savelyev et al., 2020). Reference materials such as those developed here then serve to validate the proper operation of AUC instruments. Despite the increasing prevalence of AUC in the biopharmaceutical and biotechnology industries, the U.S Food and Drug Administration (USFDA) has not recognized any molecular standards for AUC instruments (USFDA, 2023).

The precise calibration of rotor speeds, optical magnification, and rotor temperature can vary in a statistically-significant manner in AUC instruments from lab to lab. Validated molecular standards would allow for these relative calibration differences across instruments to be characterized and corrected for (Savelyev et al., 2020). This is particularly useful when developing a biopharmaceutical product that must adhere to the strict quality control standards that are mandated in the Good Manufacturing Practices (GMP) set forth by many governing bodies across the globe (Savelyev et al., 2020). Molecular Standards would aid in validating the UltraScan software's reproducibility, robustness of fitting routines, and aid in UltraScan's goal of supporting Good Manufacturing Practices (GMP) validation for the analytical ultracentrifuge (Savelyev et al., 2020). Double-stranded DNA molecules generated by recombinant means are a suitable candidate for molecular standards because their sequence is known precisely, and hence their molar masses. Furthermore, fragment lengths and anisotropy can be precisely controlled by the length of the DNA molecule, and DNA molecules are highly stable in certain buffer conditions (Matange et al., 2021).

4 Molecular Standards Production Workflow

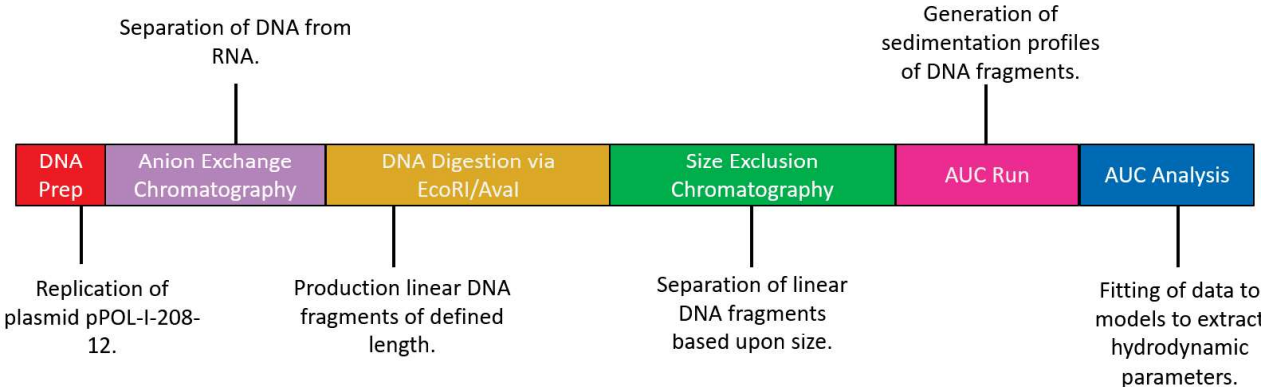

Figure 2: Diagram of the molecular standards production workflow for this project.

The goal of the project was to produce, purify, and analyze double-stranded DNA molecules intended for use as molecular standards for the analytical ultracentrifuge. This stage of the project had 6 main steps, all of which are explored in the following sections. The first was replicating the recombinant plasmid pPol1-208-12 in *E. coli* via large-volume plasmid DNA preparation. The second was using anion exchange chromatography to separate plasmid DNA from unwanted RNA fragments by exploiting the relative charge differences between the two macromolecules. The third was performing restriction endonuclease or enzyme digests using the restriction enzymes EcoRI and AvaI to produce various-sized linear fragments of defined lengths. The fourth was to separate these various sized linear fragments based upon size using size exclusion or gel

filtration chromatography. The fifth was to utilize the analytical ultracentrifuge instruments to generate sedimentation profiles of the various-sized linear DNA fragments. The final step was to fit the data collected in the analytical ultracentrifuge instrument using various techniques to extract values for various hydrodynamic parameters, such as the sedimentation coefficient, diffusion coefficient, and partial specific volume.

5 Plasmid of Interest

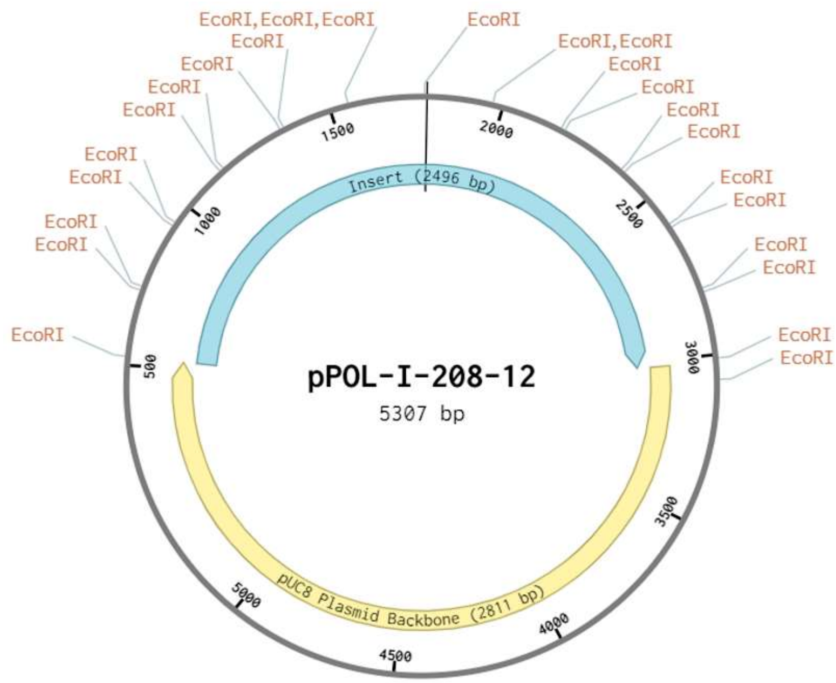

Figure 3: Map of the pPOL-I-208-12 plasmid (Martin, 2022). Comprised of a 2811 base pair pUC8 backbone and a 2496 base pair insert. The plasmid contains 25 EcoRI restriction enzyme recognition sites. When digested with EcoRI 12 x 196 base pair fragments and 12 x 12 base pair fragments were produced per plasmid.

The plasmid of interest used to develop the molecular standard is called pPOL-I-208-12. The lowercase 'p' in the abbreviation indicates that it is a plasmid, which is a small, circular, double-stranded DNA molecule that is independent of the chromosomal DNA of the organism (National Institutes of Health, 2023). They are most commonly found in bacteria, and can be used to exchange genetic advantages such as antibiotic resistance (National Institutes of Health, 2023). The 'POL-I' refers to the presence of the DNA polymerase 1 gene, which aids in DNA replication and repair (Allen et al., 2011). The '208' represents the base pair length of the DNA fragments, consisting of the 5S ribosomal RNA gene from *Xenopus laevis* that are created when cut by the AvaI restriction enzyme. The '12' represents the number of 208 base pair fragments that were produced from each plasmid when digested by the AvaI restriction enzyme. The plasmid pPol-I-208-12 is 5307 base pairs long, which consists of a 2496 base pair insert and a 2811 pUC8 plasmid backbone. The insert contains the LacZ gene which codes for the enzyme beta-galactosidase, but is deactivated in plasmids containing the insert through insertional mutation. The insert also contains the gene for ampicillin antibiotic resistance to aid in the proliferation of E. *coli* bacteria containing the plasmid of interest. With the aim of cost-effectiveness, EcoR1 restriction enzyme was used to produce 12 fragments of 196 base pairs and 12 fragments of 12 base pairs to be studied.

6 Plasmid DNA Preparation

6.1 Plasmid DNA Preparation Method

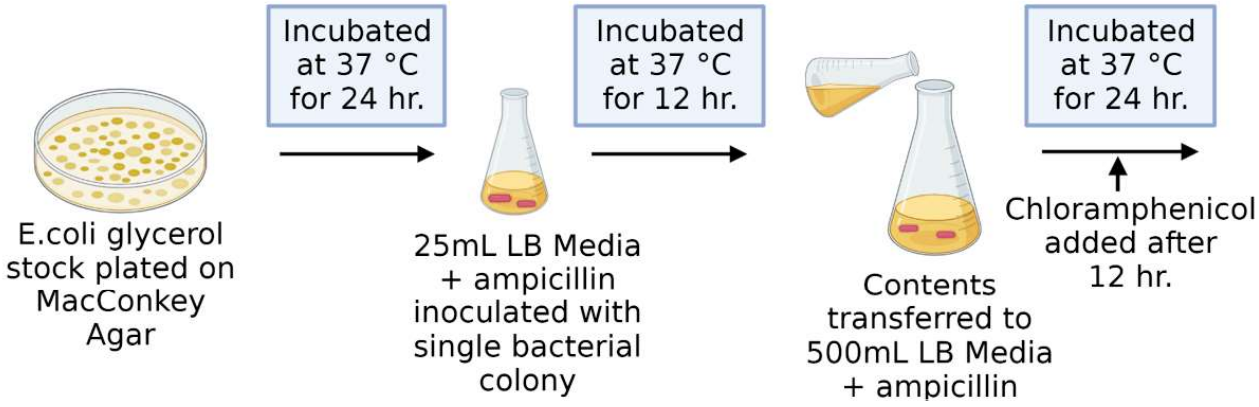

Figure 4: Flow chart depicting part of the process of a 2L plasmid DNA preparation (Martin, 2023). Initiated with glycerol stocks of *E. coli* culture, followed by various incubation stages for bacterial growth.

Figure 5: Flow chart depicting part of the process of a 2L plasmid DNA preparation (Martin, 2023). Continued with the centrifugation of the bacterial culture, treatment with various DNA extraction solutions, and resuspension of the nucleic acids in a suitable buffer.

The plasmid DNA preparation was initiated by plating E. *coli* glycerol stock on MacConkey agar plates, which were incubated at 37 degrees Celsius for 24 hours. The MacConkey plates contained a red or white test, in which the beta-galactosidase coded by the LacZ gene broke down lactic acid in the E. *coli*, and released acidic products that lowered the pH of the surrounding environment. Colonies that successfully contain the recombinant plasmid appeared white due to the lactic acid breakdown (Juers et al, 2012). After sufficient bacterial growth, a single white colony was inoculated in 25 ml of ampicillin-containing LB media for 12 hours at 37 degrees Celsius. Then, the contents were transferred into a 500 mL LB media to be incubated for 24 hours at 37 degrees Celsius.

After twelve hours, an antibiotic called chloramphenicol was added to the flasks. Chloramphenicol inhibits protein synthesis in bacteria, which prevented further growth of the bacteria so over-crowding at the end of the exponential growth phase does not occur (Amati, 1970). During overcrowding, the nutrients in the media are depleted, and the environment becomes toxic. Hence the death rate of the bacterial cells overtakes the growth rate (Amati, 1970). Chloramphenicol still allows for plasmid DNA replication to continue, so the overall yield of plasmid DNA can be increased (Amati, 1970). The bacterial culture was centrifuged and the resulting cell pellets were resuspended in DNA extraction solution I, which contained glucose to protect the

structure of the plasmid DNA (Promega Canada, 2023) . An enzyme known as lysozyme was then added to break open the bacterial cells and release their contents. Next, DNA extraction solution II was added to increase the solubility of the bacterial cell membranes (Promega Canada, 2023). Following a brief incubation period, acetate salt-containing DNA extraction solution III was added to precipitate the contents out of solution. The solution was centrifuged and the supernatant was precipitated using isopropanol alcohol. Following an additional centrifugation step, the nucleic acid-containing pellet was resuspended in a DNA-stable pH 8 Tris-EDTA (TE) buffer. The quantity of nucleic acids produced was determined using a UV-spectrophotometer in tandem with the Lambert-Beer Law.

6.2 Plasmid DNA Preparation Results

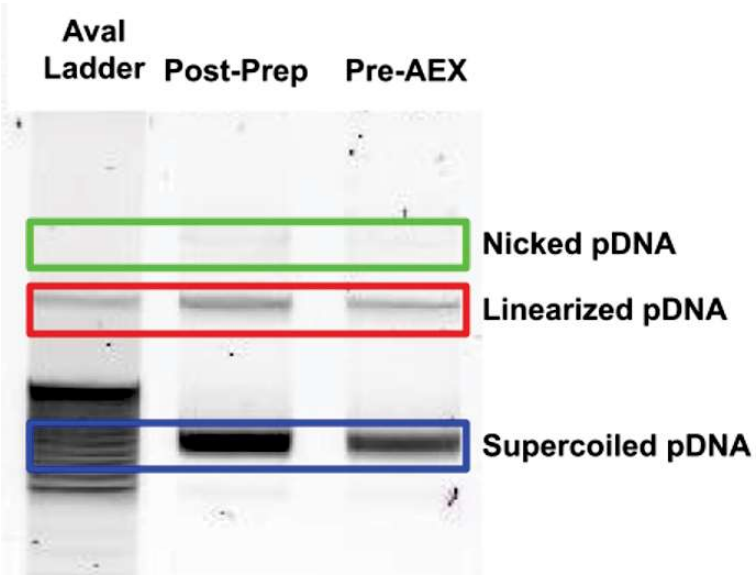

Figure 6: 1% agarose gel electrophoresis performed on the material from the plasmid DNA preparation. 0.1 µg of pDNA loaded into each well.

The aim of the plasmid DNA preparation was to produce as much of the supercoiled plasmid DNA as possible. To visualize the results, the technique of agarose gel electrophoresis was used. Gel electrophoresis separates nucleic acids based upon their size and charge [(Lee et al., 2012). In this case, the negatively-charged DNA molecules were forced to migrate through the agarose gel matrix with the application of a potential difference from the top to the bottom of the gel as it rested in an ion-containing buffer composed of Tris-HCl, acetate, and EDTA (TAE) (Lee et al., 2012). The lane labeled AvaI Ladder was used as a reference as it was a well-characterized sample with bands of known fragment length and migration rate. To separate the DNA from the contaminating RNA, a ribonuclease (RNAse) A enzyme digestion was performed on the resuspended nucleic acids sample, which degrades the RNA, as it was not of interest. To separate the nucleic acids from proteins, lipids, and any cell debris, an ethanol precipitation and microcentrifuge clarification were performed, followed by gel electrophoresis. The tightly-packed supercoiled plasmid DNA migrated the fastest as it experiences the least friction in the gel matrix. The nicked-circular plasmid DNA, where the supercoiled conformation of the plasmid DNA was broken, migrated at the slowest rate due to the greater friction it experienced in the gel matrix (Lee et al., 2012). The presence of DNA nucleases leads to the unavoidable contamination of the supercoiled plasmids with nicked and linearized pDNA (Lee et al., 2012).

7 Anion Exchange Chromatography (AEX)

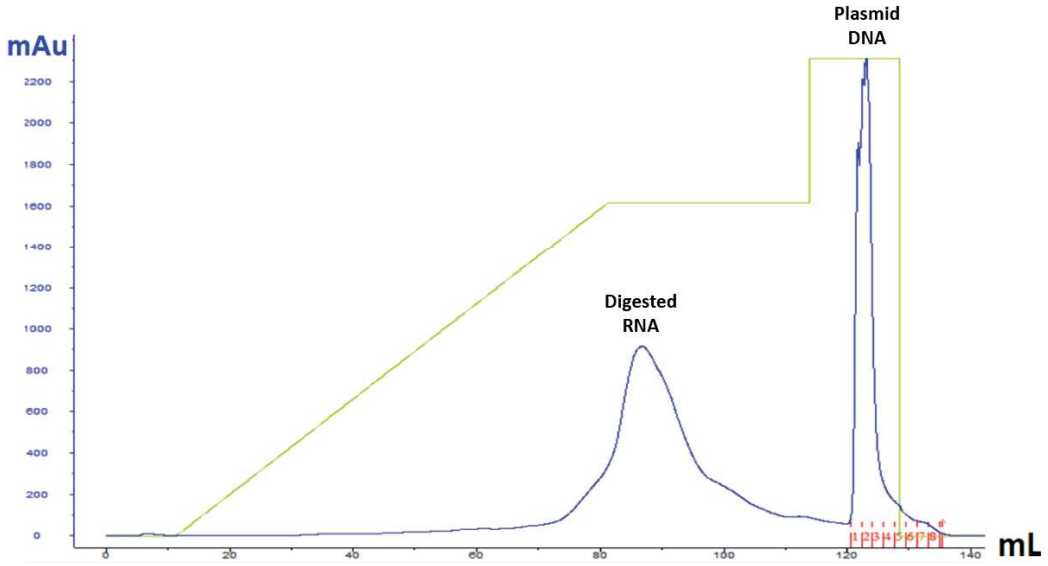

Figure 7: Anion exchange chromatogram from the FPLC fraction collector displaying absorbance (blue) in mAU and the relative NaCl concentration of the flow volume (green) as a function of the total flow volume in mL.

In order to separate the plasmid DNA from the digested RNA, the DNA preparation was purified using anion exchange (AEX) chromatography. In AEX, the negative phosphate backbone of the DNA binds to the positive functional groups in the column matrix (Budelier and Schorr, 1998). The RNA will also bind to the column due to its phosphate group. However, as DNA is more negatively charged than RNA, it is expected that RNA will elute from the column first when a steadily increasing NaCl concentration is applied (Budelier and Schorr, 1998) The hardware used to achieve this was a 5 mL HiTrap Q HP anion exchange column with an ÄKTA Prime Plus fast protein liquid chromatography (FPLC) pump to control the flow rates and produce the chromatogram using the built-in spectrophotometer. The column was first equilibrated with a 10mm Tris-HCl pH 8.06 buffer, which initially composed 100% of the flow volume.

The nucleic acids were injected onto the column, and allowed to bind to the functional groups present within the matrix. Then, in order to elute the RNA, a 10mm Tris-HCl + 1M NaCl pH 8.06 elution buffer was introduced to the flow volume on a gradient of 70% over 70 mL. In which the NaCl concentration rose from 0 mM to 700 mM at a rate of 10mM per mL of flow volume. The RNA elution was complete when the absorbance trace reached baseline on the chromatogram. At this stage, the percent of the elution buffer in the flow volume was manually increased to 100% or a concentration of 1M NaCl. The increased salt concentration caused the plasmid DNA to elute from the column in the final, sharp peak. The chromatogram above displays absorbance in mAU (absorbance units) (in blue) and the percent NaCl concentration of the total flow volume (in green) plotted as a function of the total flow volume in mL. The fraction collector separated and stored the material that flowed through the column in 1.8 mL fractions, which are indicated at bottom of the chromatogram in the red number scale. From 10 mg of nucleic acid loaded onto the column, the final, sharp peak corresponds to plasmid DNA, with a typical yield of about 500 µg of plasmid DNA.

8 Restriction Enzyme Digest

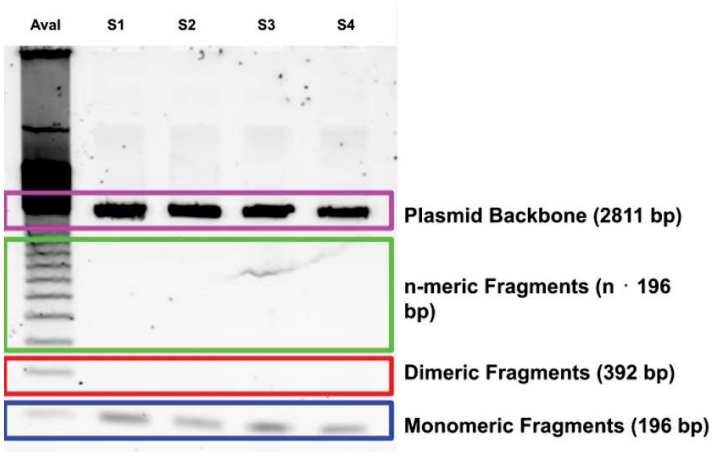

Figure 8: 1% agarose gel on an EcoRI digest of the pPOL-I-208-12 plasmid. 0.1 µg pDNA loaded per well. Samples S1-S4 were from the same reaction tube.

Once the double-stranded plasmid DNA was been purified using AEX, the DNA was digested using restriction enzymes. The restriction enzymes bind at specific sites containing the appropriate DNA recognition sequences, and cleave the strand in a similar manner to a pair of scissors (Qiagen, 2022). For Figure 8, an EcoRI restriction enzyme digestion was performed at 37 degrees Celsius for two hours using an enzyme concentration of 1.0 u/µg. These digestion parameters reliably produced a complete digest of the pPOL-I-208-12 plasmid, generating 196 base pair monomer fragments and 2811 base pair backbone fragments.

9 Size Exclusion Chromatography (SEC)

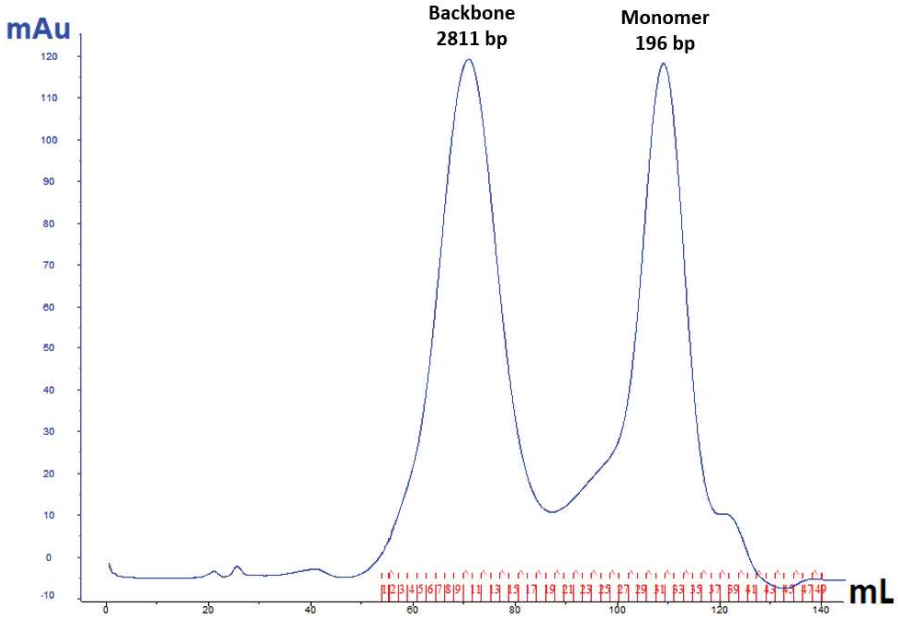

Figure 9: Chromatogram displaying the elution profile of the 2811 base pair plasmid backbone and 196 base pair monomeric DNA fragments, where absorbance (blue) in mAU is plotted as a function of elution volume in mL.

Once the mixtures of linearized plasmid DNA fragments were produced from the EcoRI restriction enzyme digests, the fragments were separated based upon size using gel filtration or size exclusion chromatography (SEC). SEC involves a vertical column that utilizes gravity flow to filter the fragments through a gel of various-sized spherical beads possessing a certain pore size (Tosoh Bioscience, 2018). Molecules that are larger than the pore size are expected to elute first due to their inability to penetrate the pores; whereas the smaller molecules take a longer time to elute as they can penetrate the pores and take an increasingly long path to reach the bottom (Tosoh Bioscience, 2018).

The goal of SEC in this case was to separate the monomeric fragments from the plasmid backbone. Greater separation of the fragments was indicated by greater separation of the two peaks displayed on the chromatogram. This can be visualized by fitting Gaussians to each of the peaks and observing the overlaps. The chromatogram above was produced by the spectrophotometry unit in an ÄKTA Prime Plus FPLC fraction collector, and plotted absorbance in mAU as a function of elution volume in mL. The column used was an XK 16/70 column containing Sephacryl S-1000 Superfine column matrix resin, which has a pore size of approximately 3000 Angstrom. The pump separates and stores the material that flows through the column in 1.8 mL fractions; which can be seen at the bottom of the chromatogram in the red number scale. The equilibration buffer used for the column was 1X TEN (10 mM Tris, 1mM EDTA, and 100 mM NaCl).

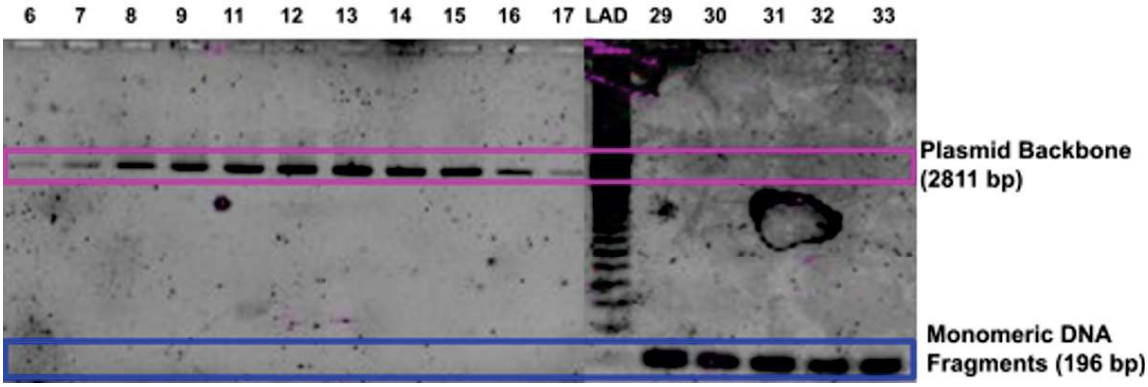

Figure 10: 1% agarose gel of selected fractions collected from the SEC peaks in Figure 9. Fractions 6-17 contained the pure plasmid backbone fragments, while fractions 29-33 contained the pure monomeric DNA fragments.

To gauge the separation of the linear DNA fragments on the column, all fractions collected from the SEC peaks in Figure 9 were run on a 1% agarose gel to observe the purity in each fraction. Fractions 6-17 were determined to contain the pure 2811 base pair backbone fragments, while fractions 29-33 were determined to contain the 196 base pair monomeric DNA fragments. Fractions 18-28 contain various mixing ratios of the fragments. Due to their lack of purity they were added to another run to be re-purified using the SEC column.

10 Analytical Ultracentrifugation Theory

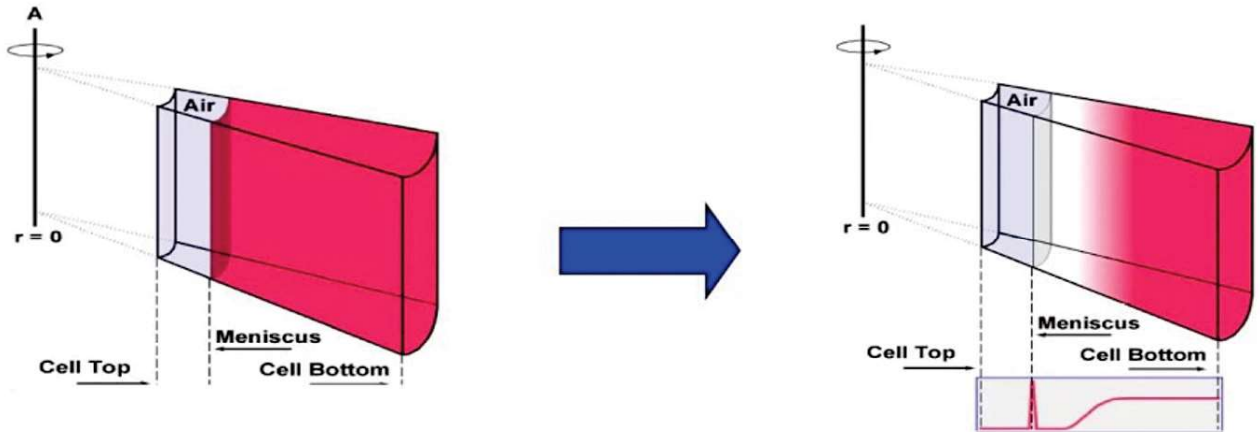

Figure 11: Time-evolution of the boundary of macromolecules as sedimentation occurs in a sector-shaped AUC cell (Zollars, 2022).

Analytical ultracentrifugation is a first-principle biophysical characterization method for biological macromolecules in the solution phase. The technique can be used to determine hydrodynamic parameters, such as the sedimentation coefficient, diffusion coefficient, partial specific volume, and purity of various analytes (Cole et al., 2008). The sedimenting behavior of the macromolecules in solution at each radial position in the sector-shaped cell are defined by the redistribution of the mass across the gravitational field until the gravitational potential energy is equivalent and opposite to the chemical potential energy. The boundary created by this redistribution, and its evolution throughout the AUC cell as the macromolecules sediment is what is monitored in defined intervals throughout the course of a sedimentation velocity AUC experiment in the form of scans (Cole et al., 2008).

$$F_{net} = (m_p - m_s)\,\omega^2 r \qquad s = \frac{v}{\omega^2 r} = \frac{m_p\,(1 - \overline{v}\rho)}{f}$$

$$F_f = -fv \qquad \text{Svedberg Equation: } \frac{s}{D} = \frac{M_p\,(1 - \overline{v}\rho)}{RT}$$

$$D = \frac{RT}{N_a f}$$

Lamm Equation:

$$\frac{\partial c}{\partial t} = D\left[\frac{\partial^2 c}{\partial r^2} + \frac{1}{r}\frac{\partial c}{\partial r}\right] - s\omega^2\left[r\frac{\partial c}{\partial r} + 2c\right]$$

Figure 12: Various equations describing the forces acting within an AUC cell, as well as the relevant hydrodynamic parameters and the equations used to extract values for them (Cole et al., 2008).

The net force acting on the sedimenting particle is given by the difference between the mass of the particle m_p, and the mass of the solvent m_s displaced; multiplied by the rotor speed ω in radians per second, and the radial distance r from the center of the rotor (Cole et al., 2008). The force due to friction acting on the particle is given by the velocity v, and frictional coefficient f of the particle. The diffusion coefficient D was derived from the Stokes-Einstein relation, where R is the gas constant, T is the absolute temperature, and N_a is Avogadro's number (Cole et al., 2008). The sedimentation coefficient s is the sedimentation rate of the particle reported in units of Svedbergs, or 10^{-13} seconds. $\bar{v}$ is the partial specific volume (reciprocal of the density) of the particle, and ρ is the density of the solvent (Cole et al., 2008). The sedimentation coefficient depends on the size and shape of the molecule, so larger molecules will possess a larger sedimentation coefficient value.

The Svedberg equation or the ratio of the sedimentation and diffusion coefficients, where M_p is the molar mass of the sedimenting particle. For the Lamm equation for interacting systems, where c is the concentration of the solute along the boundary (Cole et al., 2008). The equation describes the evolution of the radial concentration distribution in the AUC cell over time for the sedimenting particles. Solutions to this partial differential equation cannot be determined analytically, so numerical solutions must be attempted using the finite element method (Schuck and Zhao, 2017). In tandem, the Svedberg and Lamm equations can be fitted to extract the highlighted hydrodynamic parameters and understand some of the physical properties of the sedimenting macromolecules (Cole et al., 2008).

11 Analytical Ultracentrifuge Results

11.1 Analytical Ultracentrifuge Data

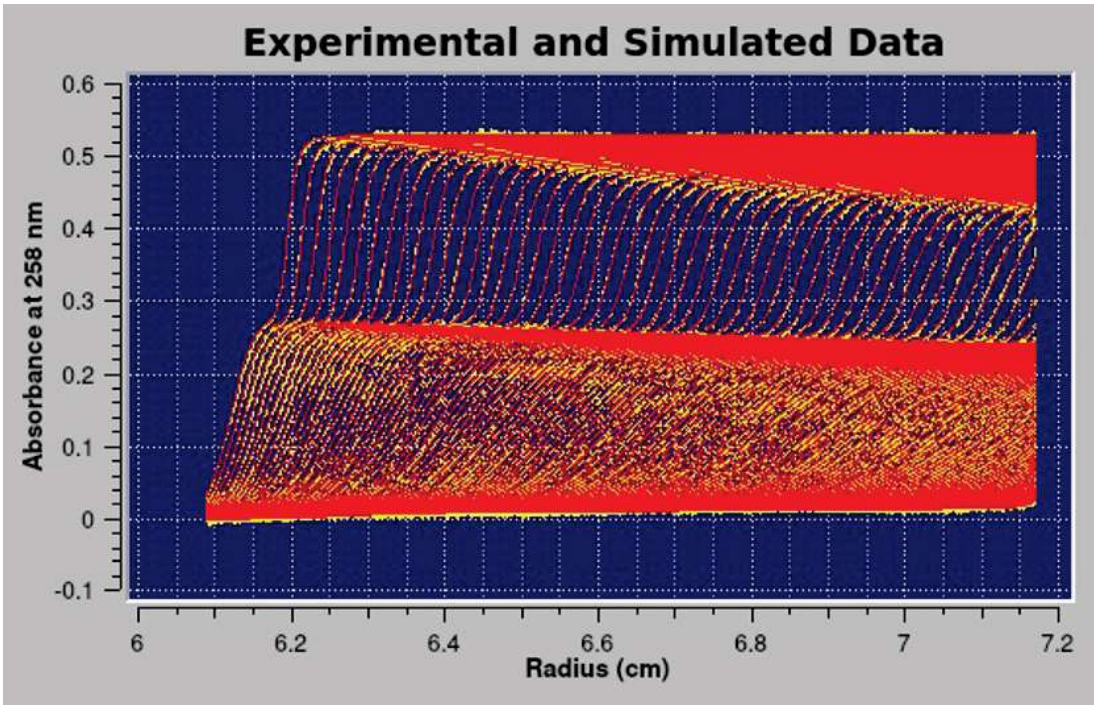

Figure 13: Sedimentation profile of an equal mixing ratio of the 196 base pair monomeric DNA fragments, and the 2811 base pair plasmid backbone fragments displayed in Figure 10. The absorbance at 258 nm is plotted as a function of the radial position in the cell. Raw AUC data is presented as yellow traces, while the modelled data is presented as red traces (Data collected by Sophia Bird[2])

A sedimentation velocity AUC experiment was then performed to develop a sedimentation profile of the linear DNA fragments from Figure 10. The UV/Vis absorbance optical system was used to exploit the light-absorbing properties of DNA, which primarily absorbs light around a wavelength of 260 nm (Lee, 2017). The results were then imported and analyzed using the UltraScan software package. Figure 13 shows the sedimentation velocity profile of a sample consisting of equal mixing ratios of the 196 base pair monomeric DNA fragments, and the 2811 base pair plasmid backbone fragments displayed in Figure 10. Absorbance at 258 nm is plotted as a function of the radial position in the cell. The traces in red correspond to the modelled data, while the yellow traces correspond to the raw data collected in the experiment. The agreement between the experimental and simulated data is given by the root mean squared deviation (RMSD) between the two sets. In this case, the RMSD value was 0.0025, which indicated that the model provided a very good fit of the data. The specific parameters of this run included a rotor speed of 35,000 RPM, a temperature of 20 °C, and a run time of 12 hours. The buffer used for this experiment was a 50mM Tris, 1mM EDTA, 150 mM NaCl, and 20 mM $CaCl_2$. (Data collected by Sophia Bird[2]).

11.2 Analytical Ultracentrifuge Data Analysis

11.2.1 Sedimentation and Diffusion Coefficient Distribution

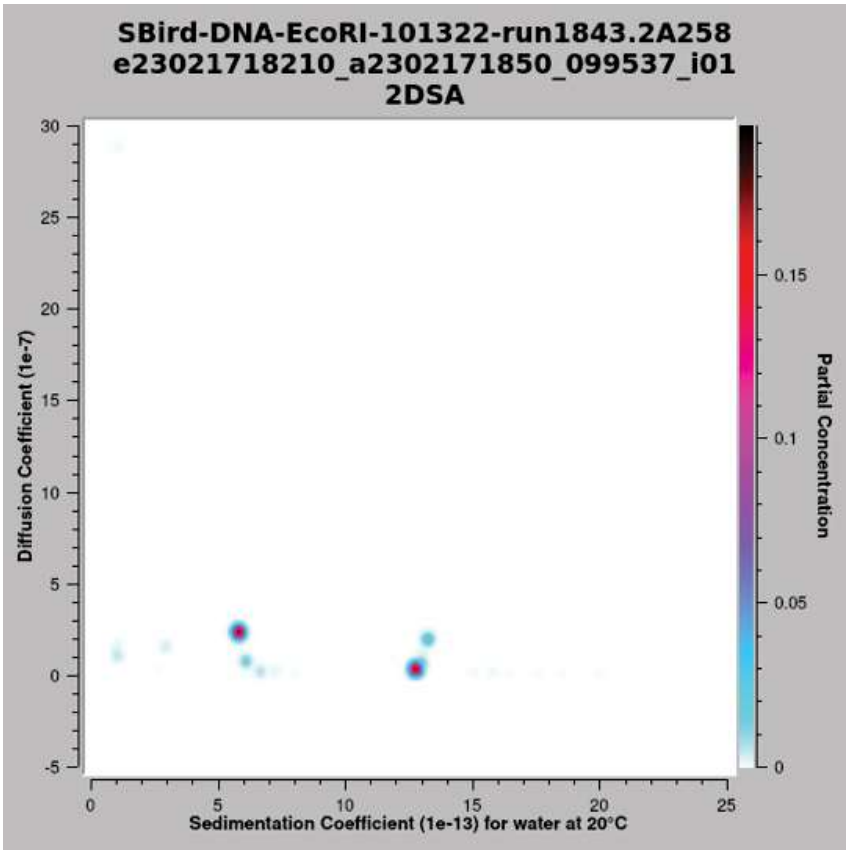

Figure 14: Diffusion and sedimentation coefficient distributions as a function of the analyte partial concentration. Figure 14 was generated after one 2DSA iteration. Larger molecules, such as the 2811 base pair plasmid backbone are expected to have a higher sedimentation coefficient than the smaller 196 base pair monomeric fragments. (Data collected by Sophia Bird[2])

After performing one iteration of a two-dimensional spectrum analysis (2DSA) on the experimental data set, and eliminating time-invariant noise, diffusion and sedimentation coefficient distributions as a function of the analyte partial concentration were produced (Walter et al., 2016). Larger molecules such as the 2811 base pair plasmid backbone were expected to have a higher sedimentation coefficient than the smaller 196 base pair monomeric fragments. The apparent species from this distribution were then compared to the expected molecular weights. In this case, the expected molecular weight of the 196 base pair monomeric fragments was 121,136.54 Da. After performing a Monte Carlo analysis, the species around 6 Svedberg was determined to have an average molecular weight of about 121,100 Da. This suggested that this species was likely composed of the 196 base pair monomeric fragments. The expected molecular weight of the 2811 base pair backbone fragment is 1,736,810.95 Da. The species around 13 Svedberg was determined to have an average molecular weight of about 1,736,800 Da. This suggested that this species was likely composed of the 2811 base pair plasmid backbone fragments.

11.2.2 Discrete Sedimentation Coefficient Distribution

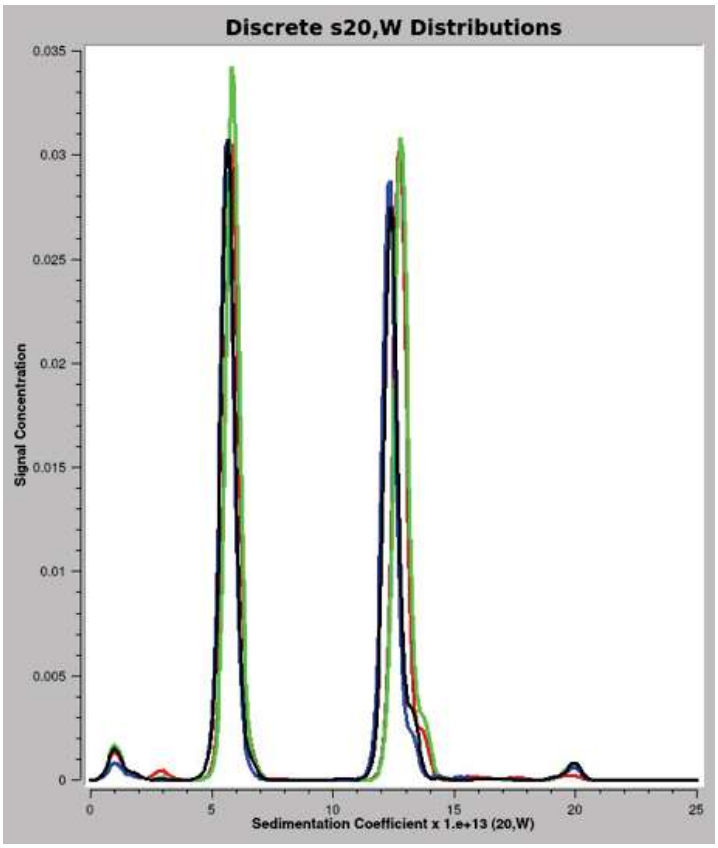

Figure 15: Relative signal concentration of various-sized linear DNA fragments. Larger molecules such as the 2811 base pair plasmid backbone are expected to have a higher sedimentation coefficient than smaller molecules such as the 196 base pair monomeric fragments. (Data collected by Sophia Bird[2]).

SBird-DNA-EcoRI-101322-run1843 (2A258, 2DSA-MC) pPOL1-208-12 in 10 mM TRIS 20 mM CaCl2 150 mM NaCl[14])
SBird-DNA-EcoRI-101322-run1843 (2B258, 2DSA-MC) pPOL1-208-12 in 10 mM TRIS 40 mM CaCl2 150 mM NaCl[29])
SBird-DNA-EcoRI-101322-run1843 (4A258, 2DSA-MC) pPOL1-208-12 in 10 mM TRIS 20 mM CaCl2 500 mM NaCl[44])
SBird-DNA-EcoRI-101322-run1843 (4B258, 2DSA-MC) pPOL1-208-12 in 10 mM TRIS 40 mM CaCl2 500 mM NaCl[59])

Figure 16: Legend for Figure 15, which corresponds each curve to the appropriate triple data set from the sedimentation velocity experiment.

The discrete sedimentation coefficient distribution in Figure 15 plotted a distribution of the relative signal concentration of the analytes as a function of the sedimentation coefficient (Walter et al., 2016). Larger molecules such as the 2811 base pair plasmid backbone (~ 13 *s*) were measured to have a higher sedimentation coefficient than smaller molecules such as the 196 base pair monomeric fragments (~ 6 *s*). The smaller species around 1 *s* were likely the 12 base pair fragments, and the larger species around 20 *s* were likely undigested plasmid.

12 DNA Fragment Production Optimization

12.1 Restriction Digestion Optimization

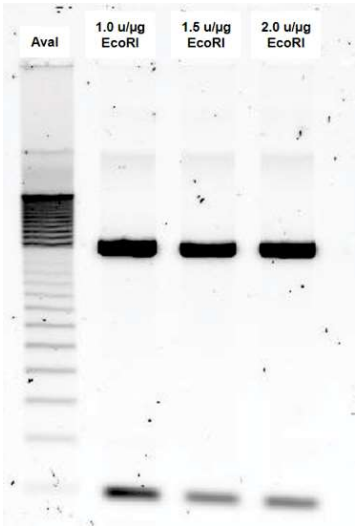

Figure 17: 1% agarose gel on a 2-hour EcoRI digestion of pDNA with various enzyme concentrations.

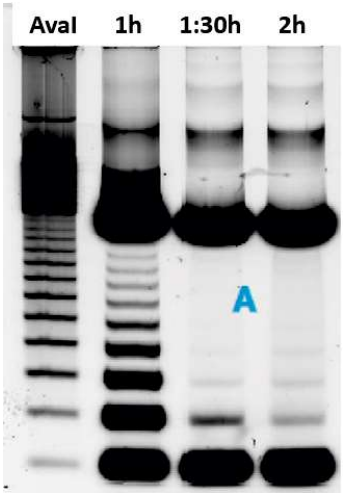

Figure 18: 1% agarose gel on a 1.0 u/µg EcoRI digestion of pDNA with various incubation times.

One of the objectives set out for the Fall 2022 semester was to optimize the EcoRI restriction enzyme digests, both in terms of the restriction enzyme concentration required, and the incubation times for the digest. Figure 17 demonstrated that increasing EcoRI restriction enzyme concentration did not noticeably affect the digestion profiles after two hours. Figure 18 demonstrated that incubation times between one and two hours had a significant effect on the composition of the resulting digestion profiles. This is indicated by the lack of complete digestion of the plasmid DNA when the incubation time is reduced. The takeaway from this was that incubation time has a greater impact on the digestion profiles than restriction enzyme concentration.

12.2 Size Exclusion Chromatography Optimization

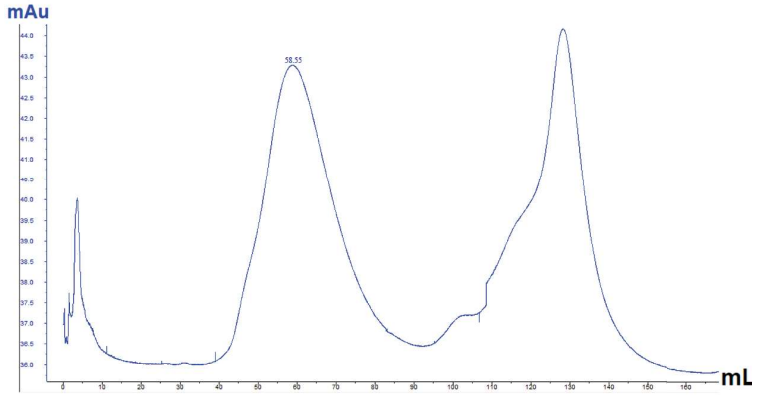

Figure 19: SEC chromatogram. 0.5 mg pDNA in 0.5 mL. pDNA load concentration is 1 mg/mL.

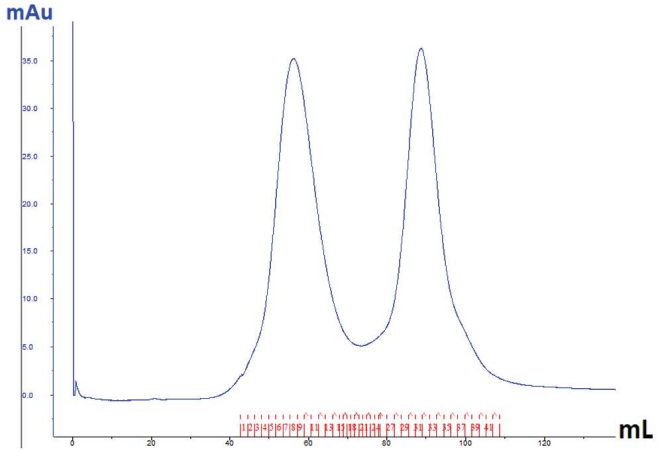

Figure 20: SEC chromatogram. 0.5 mg pDNA in 1 mL. pDNA load concentration is 0.5 mg/mL.

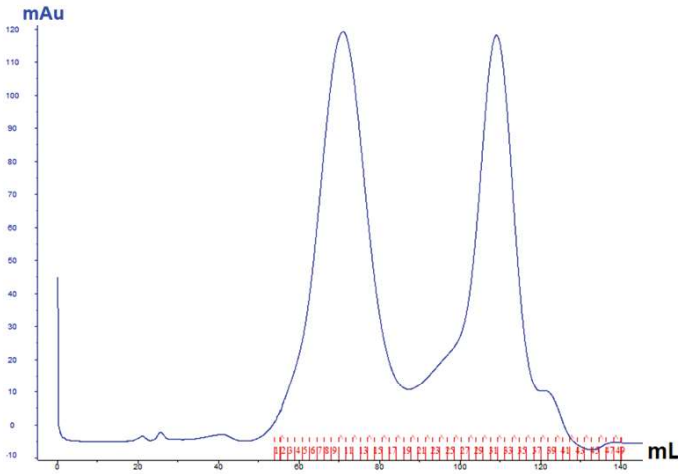

Figure 21: SEC chromatogram. pDNA load concentration 0.5 mg/mL.The resin bed height was extended from 61.5 cm in Figure 20 to 67 cm for this Figure.

Another objective set out for the Fall 2022 semester was to optimize the peak resolution and peak separation of the size-exclusion column. Peak resolution was optimized by lowering the pDNA load concentration from 1 mg/mL in Figure 19 to 0.5 mg/mL in Figure 20. In Figure 20, the peaks appear sharper compared to Figure 19. From size-exclusion chromatography theory, using a longer resin bed in the column increases the available pore volume and therefore increases peak separation (Tosoh Bioscience, 2018). So peak separation was optimized by extending the resin bed height from 61.5 cm in Figure 20 to 67 cm in Figure 21. In Figure 21, the peaks appear better separated compared to Figure 20.

13 Conclusion and Future Directions

First, the replication of the pPOL-I-208-12 plasmid in E. *coli* via a large-volume plasmid DNA preparation was optimized for high yields. Next, the purification of the nucleic acids via AEX was optimized for large yields of supercoiled plasmid DNA with no detectable contamination from unwanted RNA fragments. After this, the production of the 196 base pair monomeric fragments and 2811 base pair backbone fragments via EcoRI restriction enzyme digests were optimized both in terms of digestion time and restriction enzyme concentration. Next, the separation of the various-sized plasmid DNA fragments via SEC was mostly optimized for efficient purification. Finally, it was demonstrated that the 50mM Tris, 1mM EDTA, 150 mM NaCl, and 20 mM $CaCl_2$ buffer produced acceptable conditions for running the monomeric and backbone fragments in the AUC instrument. This was evidenced by the minimized concentration-dependent non-ideality behaviour, and the acceptable RMSD-values of the fitted data. Moving forward, a goal is to further extend the resin bed height in the size-exclusion column to achieve better peak separation, particularly as attempts are made to isolate higher-order n-meric fragments. Another goal is to extensively characterize the monomer and backbone fragments in the analytical ultracentrifuge under varying temperatures, salt concentrations, and rotor speeds. After this, the focus will move on to characterizing the higher-order n-meric fragments in the same manner.

I would like to thank the following individuals and organizations for making this work possible: Dr Borries Demeler[2,3], Dr Ken Vos[1], Sophia Bird[2], Floriane Houenagnon[2], Amy Henrickson[2], Owen Woycenko, Thea Demeler (MITACS), Agraj Paudel (HYRS), Dryden Many Bears (InSRA), the Canadian Center for Hydrodynamics (CCH), Natural Sciences and Engineering Research Council of Canada (NSERC), and the NSERC RNA Innovation program.

[1]Department of Physics and Astronomy, University of Lethbridge, Alberta, Canada, T1K 3M4.
[2]Department of Chemistry and Biochemistry, University of Lethbridge, Alberta, Canada, T1K 3M4.
[3]Department of Chemistry and Biochemistry, University of Montana, Missoula, MT, USA, 59812.

14 References

Allen, J. M., Simcha, D. M., Ericson, N. G., Alexander, D. L., Marquette, J. T., Van Biber, B. P., Troll, C. J., Karchin, R., Bielas, J. H., Loeb, L. A., & Camps, M. (2011). Roles of DNA polymerase I in leading and lagging-strand replication defined by a high-resolution mutation footprint of ColE1 plasmid replication. *Nucleic Acids Research*, 39(16), 7020–7033. https://doi.org/10.1093/nar/gkr157

Amati P. (1970). Chloramphenicol: Effect on DNA synthesis during phage development in Escherichia *coli*. *Science* (New York, N.Y.), 168(3936), 1226–1228. https://doi.org/10.1126/science.168.3936.1226

Budelier, K. and Schorr, J. (1998), Purification of DNA by Anion-Exchange Chromatography. *Current Protocols in Molecular Biology*, 42: 2.1.11-2.1.18. https://doi.org/10.1002/0471142727.mb0201bs42

Cole, J. L., Lary, J. W., P Moody, T., & Laue, T. M. (2008). Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. *Methods in Cell Biology*, 84, 143–179. https://doi.org/10.1016/S0091-679X(07)84006-4

Juers, D. H., Matthews, B. W., & Huber, R. E. (2012). LacZ β-galactosidase: structure and function of an enzyme of historical and molecular biological importance. *Protein Science*: A publication of the Protein Society, 21(12), 1792–1807. https://doi.org/10.1002/pro.2165

Lee, A. (2017). DNA Concentration Measurement at 260 nm. *Tip Biosystems*. Retrieved from https://www.semanticscholar.org/paper/DNA-CONCENTRATION-MEASUREMENT-AT-260-nm-USING-BIO-Lee/a72668c4f60a046293fd3135c75b0823926df6d8

Lee, P. Y., Costumbrado, J., Hsu, C. Y., & Kim, Y. H. (2012). Agarose gel electrophoresis for the separation of DNA fragments. *Journal of Visualized Experiments (JoVE)*, (62), 3923. https://doi.org/10.3791/3923

Martin, R. (2022, May 12). *pPOL-I-208-12 - Benchling Editor*. Benchling. Retrieved from https://benchling.com/rks2000/f/lib_4IRgF65Q-ppol-i-208-12/seq_vReEhF5J-ppol-i-208-12/edit

Martin, R. (2023, March 12). *Plasmid DNA Replication Diagram*. BioRender. Retrieved from https://app.biorender.com/illustrations/6320f5e921a57b6b945b0b21

Matange, K., Tuck, J. M., & Keung, A. J. (2021). DNA Stability: A central design consideration for DNA data storage systems. *Nature Communications*, 12(1), 1358. https://doi.org/10.1038/s41467-021-21587-5

National Institutes of Health. (2023, April 18). *Plasmids*. Genome.gov . Retrieved from https://www.genome.gov/genetics-glossary/Plasmid#:~:text=A%5C%20plasmid%5C%20is%5C%20a%5C%20small,chromosomal%5C%20DNA%5C%20and%5C%20replicate%5C%20independently.

Promega Canada. (2021). *DNA Purification*. Retrieved from https://www.promega.ca/resources/guides/nucleic-acid-analysis/dna-purification/

QIAGEN. (2022). *Restriction Endonuclease Digestion of DNA*. QIAGEN. Retrieved from https://www.qiagen.com/us/knowledge-and-support/knowledge-hub/bench-guide/dna/handling-dna/restriction-endonuclease-digestion-of-dna#:~:text=Many%20applications%20require%20conversion%20of,DNA%20at%20specific%20target%20sequences.

Savelyev, A., Gorbet, G. E., Henrickson, A., & Demeler, B. (2020). Moving analytical ultracentrifugation software to a good manufacturing practices (GMP) environment. *PLoS Computational Biology*, 16(6), e1007942. https://doi.org/10.1371/journal.pcbi.1007942

Schuck, P., and Zhao, H. (2017). *Sedimentation Velocity Analytical Ultracentrifugation: Interacting Systems* (1st ed.). pp. 1-271. CRC Press. ISBN: 9781315268705. https://doi.org/10.1201/b21988

Tosoh Bioscience. (2018). *Size Exclusion Chromatography - Columns*. Tosoh Bioscience. Retrieved from https://www.separations.eu.tosohbioscience.com/products/hplc-columns-uhplc-columns/size-exclusion

USFDA. (2023). *Recognized Consensus Standards*. United States Food and Drug Administration. Retrieved from https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfstandards/results.cfm

Walter, J., Gorbet, G., Akdas, T., Segets, D., Demeler, B., & Peukert, W. (2016). 2D analysis of polydisperse core-shell nanoparticles using analytical ultracentrifugation. *The Analyst*, 142(1), 206–217. https://doi.org/10.1039/c6an02236g

Zollars, D. (2022). *Sedimentation Equilibrium Experiments*. UltraScan Analysis Software – AUC Solutions. Retrieved from https://resources.aucsolutions.com/equilibrium.php